

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:21:26 ; Search time 107.065 Seconds
(without alignments)
4561.330 Million cell updates/sec

Title: US-09-973-994-131
Perfect score: 261
Sequence: 1 aggtgaccgtacagcattta.....gggtttttccacgtcacct 261

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.6	14.4	144362	3	US-09-949-016-16066 Sequence 16066, A
2	36.8	14.1	357	3	US-09-640-211A-1581 Sequence 1581, Ap
C 3	35.8	13.7	264206	3	US-09-949-016-12731 Sequence 12731, A
C 4	35.8	13.7	264304	3	US-09-949-016-13249 Sequence 13249, A
C 5	35.6	13.6	117391	3	US-09-949-016-13945 Sequence 13945, A
6	34.8	13.3	139049	3	US-09-949-016-17030 Sequence 17030, A
7	34.8	13.3	193555	3	US-09-949-016-15553 Sequence 15553, A
8	34.8	13.3	193555	3	US-09-949-016-15554 Sequence 15554, A
9	34.8	13.3	193555	3	US-09-949-016-15555 Sequence 15555, A
10	34.6	13.3	200663	3	US-09-949-016-12569 Sequence 12569, A
11	34.6	13.3	203093	3	US-09-949-016-14445 Sequence 14445, A
C 12	33.6	12.9	601	3	US-09-949-016-53593 Sequence 53593, A
13	33.6	12.9	57392	3	US-09-949-016-12070 Sequence 12070, A
14	33.6	12.9	57402	3	US-09-949-016-13293 Sequence 13293, A
C 15	33.4	12.8	786	3	US-09-134-000C-905 Sequence 905, App
C 16	33.2	12.7	43360	3	US-09-453-702B-206 Sequence 206, App
C 17	33.2	12.7	43360	3	US-10-114-170-206 Sequence 206, App
C 18	33.2	12.7	45325	3	US-09-453-702B-261 Sequence 261, App
C 19	33.2	12.7	45325	3	US-10-114-170-261 Sequence 261, App
20	33.2	12.7	156324	3	US-09-949-016-13749 Sequence 13749, A
C 21	33	12.6	237863	3	US-09-949-016-13404 Sequence 13404, A
22	32.8	12.6	150	5	US-09-974-300-7683 Sequence 7683, Ap
C 23	32.8	12.6	56374	3	US-09-949-002-645 Sequence 645, App

C 24	32.8	12.6	56375	3	US-09-949-002-774 Sequence 774, App
C 25	32.8	12.6	202001	3	US-09-734-674-3 Sequence 3, Appli
C 26	32.8	12.6	202001	3	US-10-274-990-3 Sequence 3, Appli
C 27	32.6	12.5	56976	3	US-09-949-016-17486 Sequence 17486, A
28	32.6	12.5	98567	3	US-09-949-016-11750 Sequence 11750, A
29	32.6	12.5	100567	3	US-09-949-016-16934 Sequence 16934, A
C 30	32.6	12.5	101674	3	US-09-949-016-12033 Sequence 12033, A
C 31	32.4	12.4	601	3	US-09-949-016-113225 Sequence 113225, A
32	32.4	12.4	285986	3	US-09-949-016-12287 Sequence 12287, A
33	32.4	12.4	288031	3	US-09-949-016-14864 Sequence 14864, A
C 34	32.4	12.4	340380	3	US-09-949-016-14179 Sequence 14179, A
35	32.2	12.3	436	3	US-09-513-999C-30681 Sequence 30681, A
36	32.2	12.3	1709	3	US-09-270-767-14410 Sequence 14410, A
37	32	12.3	39982	3	US-09-820-924-3 Sequence 3, Appli
38	32	12.3	39982	3	US-10-369-626-3 Sequence 3, Appli
39	32	12.3	51043	3	US-09-949-016-12739 Sequence 12739, A
40	32	12.3	51046	3	US-09-949-016-13946 Sequence 13946, A
41	32	12.3	144158	3	US-09-949-016-11755 Sequence 11755, A
42	32	12.3	144158	3	US-09-949-016-12936 Sequence 12936, A
C 43	31.8	12.2	1664976	3	US-08-916-421B-1 Sequence 1, Appli
C 44	31.8	12.2	1664976	3	US-09-692-570-1 Sequence 1, Appli
45	31.6	12.1	366	3	US-08-956-171E-2607 Sequence 2607, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-16066/c
; Sequence 16066, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16066
; LENGTH: 144362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(144362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16066

Query Match	14.4%	Score	37.6;	DB 3;	Length	144362;
Best Local Similarity	53.4%	Pred.	No. 1.1;			
Matches	79;	Conservative	0;	Mismatches	69;	Indels 0; Gaps 0;
QY	92	AAAGTTGCTCGCTGATTGAGCAGGAAACGAGATAAGTTGTAAATTTTGTCTCGCT	151			
Db	105133	ACAGGTGAATTGATTATGTAAACCACTAATATGGAACGAGAAATATATATCTCT	105074			
QY	152	GATTTTGTGAATATTTCTCTCACTATAAAAAGCATTTTCCAGAAATAAGAAGGAGCT	211			
Db	105073	ATTGTTAGGTTAGCAAAATTAATCTTACAATAAAAATCAGTTATCAATCATTAAGGGCTT	105014			
QY	212	TTCGAACTGGTTTTCCTCCCAAGAGTTGTA	239			
Db	105013	TTCAAAAAGCATGTCAAAGAGCAGTA	104986			

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:46:32 ; Search time 779.324 Seconds
(without alignments)
4115.199 Million cell updates/sec

Title: US-09-973-994-131
Perfect score: 261
Sequence: 1 aggtgaccgtacagcattta.....gggtttttccacgtcacct 261

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	100.0	261	10	US-10-651-991-131
2	37.4	14.3	40324	8	US-10-433-793-180
3	37	14.2	346	4	US-09-925-065A-125478
4	37	14.2	346	4	US-09-925-065A-125480
5	37	14.2	346	4	US-09-925-065A-125481
6	37	14.2	346	5	US-09-925-065A-125478
7	37	14.2	346	5	US-09-925-065A-125480
8	37	14.2	346	5	US-09-925-065A-125481
9	37	14.2	558	4	US-09-925-065A-439656
10	37	14.2	558	5	US-09-925-065A-439656
11	36.8	14.1	357	9	US-10-856-499-1581
12	36.8	14.1	5276	7	US-10-311-455-800
13	36.8	14.1	5276	8	US-10-221-714A-102
14	36.8	14.1	40862	7	US-10-311-455-2046
15	36.6	14.0	346	4	US-09-925-065A-125479
16	36.6	14.0	346	5	US-09-925-065A-125479
17	36	13.8	37515	8	US-10-433-793-27

C 18	35.8	13.7	608	4	US-09-925-065A-787534	Sequence 787534,
C 19	35.8	13.7	608	5	US-09-925-065A-787534	Sequence 787534,
20	35.8	13.7	611	4	US-09-925-065A-73437	Sequence 73437, A
21	35.8	13.7	611	5	US-09-925-065A-73437	Sequence 73437, A
22	35.8	13.7	611	12	US-10-301-480-174676	Sequence 174676,
23	35.8	13.7	611	12	US-10-301-480-788085	Sequence 788085,
24	35.8	13.7	654	4	US-09-925-065A-778843	Sequence 778843,
25	35.8	13.7	654	4	US-09-925-065A-778844	Sequence 778844,
26	35.8	13.7	654	5	US-09-925-065A-778843	Sequence 778843,
27	35.8	13.7	654	5	US-09-925-065A-778844	Sequence 778844,
28	35.8	13.7	680	4	US-09-925-065A-785063	Sequence 785063,
29	35.8	13.7	680	4	US-09-925-065A-785064	Sequence 785064,
30	35.8	13.7	680	4	US-09-925-065A-785065	Sequence 785065,
31	35.8	13.7	680	5	US-09-925-065A-785063	Sequence 785063,
32	35.8	13.7	680	5	US-09-925-065A-785064	Sequence 785064,
33	35.8	13.7	680	5	US-09-925-065A-785065	Sequence 785065,
34	35.8	13.7	823	6	US-10-027-632-172361	Sequence 172361,
C 35	35.8	13.7	823	7	US-10-027-632-172361	Sequence 172361,
36	35.8	13.7	2325	6	US-10-027-632-103528	Sequence 103528,
37	35.8	13.7	2325	6	US-10-027-632-103529	Sequence 103529,
38	35.8	13.7	2325	7	US-10-027-632-103528	Sequence 103528,
39	35.8	13.7	2325	7	US-10-027-632-103529	Sequence 103529,
C 40	35.8	13.7	561515	8	US-10-741-601-5682	Sequence 5682, Ap
C 41	35.8	13.7	561515	9	US-10-741-600-17730	Sequence 17730, A
42	35.8	13.7	3673778	7	US-10-312-841-2	Sequence 2, Appli
C 43	35.4	13.6	969	12	US-10-301-480-570600	Sequence 570600,
C 44	35.4	13.6	969	12	US-10-301-480-1184009	Sequence 1184009,
C 45	35.4	13.6	987	12	US-10-301-480-603635	Sequence 603635,

ALIGNMENTS

RESULT 1

US-10-651-991-131
; Sequence 131, Application US/10651991
; Publication No. US20050125161A1
; GENERAL INFORMATION:
; APPLICANT: CAIRNEY, JOHN
; APPLICANT: XU, NANFIE
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED CONIFER CDNAS, AND THEIR USE
; TITLE OF INVENTION: IN IMPROVING SOMATIC EMBRYOGENESIS
; FILE REFERENCE: 7648.0023-00
; CURRENT APPLICATION NUMBER: US/10/651.991
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: 60/239,250
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/260,882
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 339
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 131
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Pinus taeda
US-10-651-991-131

QY	1	AGGTGACCGTACAGCATTATTTGATTTCTATTTTGTGTTTGCAGTTTTCGATTTCG 60	Query Match	100.0%;	Score 261;	DB 10;	Length 261;
Db	1	AGGTGACCGTACAGCATTATTTGATTTCTATTTTGTGTTTGCAGTTTTCGATTTCG 60	Best Local Similarity	100.0%;	Pred. No. 1.3e-63;	Mismatches 0;	Indels 0; Gaps 0;
QY	61	CTGTGAGGCACGGAAACGAGATAAGTTGTAAAGTTTGTCTCGCTGATTGAGGCACGGA 120	Matches 261;	Conservative 0;			
Db	61	CTGTGAGGCACGGAAACGAGATAAGTTGTAAAGTTTGTCTCGCTGATTGAGGCACGGA 120					
QY	121	AAACGAGATAAGTTGTAAATTTTGTCTCGCTGATTTTGTCTGAATATTTCTCTCACTAT 180					
Db	121	AAACGAGATAAGTTGTAAATTTTGTCTCGCTGATTTTGTCTGAATATTTCTCTCACTAT 180					

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:01:45 ; Search time 333.602 Seconds
(without alignments)
5454.872 Million cell updates/sec

Title: US-09-973-994-131
Perfect score: 261
Sequence: 1 aggtgaccgtacagcattta.....gggtttttccacggtcacct 261

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_8:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	261	100.0	261	14 AEA49636	Aea49636 Loblolly
2	37.4	14.3	40324	6 ABQ67150	Abq67150 Human ang
3	36.8	14.1	357	3 AAC56859	Aac56859 Pinus rad
4	36.8	14.1	1943	14 ADW16980	Adw16980 Pinus rad
5	36.8	14.1	5276	4 AAS46380	Aas46380 Tumour su
6	36.8	14.1	5276	6 ABL32827	Ab132827 Human imm
7	36.8	14.1	40862	6 ABL34073	Ab134073 Human imm
8	36.4	13.9	1084	3 AAC59973	Aac59973 Human sec
9	36	13.8	37515	6 ABQ66997	Abq66997 Human ang
C 10	35.4	13.6	2000	11 ACL37108	ACL37108 Rice stre
11	35.4	13.6	6113	6 ABL32803	Ab132803 Human imm
C 12	35	13.4	1726	11 ADO77946	Ado77946 Brassica
13	35	13.4	73334	6 ABL34124	Ab134124 Human imm
14	35	13.4	73334	6 ABL92318	Ab192318 Chemicall
15	34.8	13.3	5542	6 ABL34021	Ab134021 Human imm
16	34.8	13.3	11209	6 ABN80136	Abn80136 Human che
17	34.6	13.3	2441	14 ADW10461	Adw10461 Colon pro
18	34.6	13.3	7135	4 AAS46423	Aas46423 Tumour su

C 19	34.4	13.2	857	4 AAI94399	Aai94399 Human neu
C 20	34.4	13.2	5272	13 ADR07576	Adr07576 Full leng
21	34.4	13.2	110000	2 AAX20248_06	Continuation (7 of
22	34.4	13.2	111309	2 AAX20250	Aax20250 Borrelia
23	34.2	13.1	353	4 AAL11902	Aal11902 Human bre
24	34.2	13.1	362	4 AAL20792	Aal20792 Human bre
25	34.2	13.1	391	4 AAL11979	Aal11979 Human bre
26	34.2	13.1	424	11 ACN82095	Acn82095 Breast ca
C 27	34.2	13.1	505	11 ACN90516	Acn90516 Breast ca
28	34.2	13.1	888	11 ACN82168	Acn82168 Breast ca
29	34.2	13.1	18683	6 ABL32312	Ab132312 Human imm
30	34.2	13.1	18683	6 ABL54333	Ab154333 Chemicall
31	34	13.0	3025	8 ABZ10190	Abz10190 Haematopo
32	34	13.0	3025	10 ADB54220	Adb54220 Pretreat
33	34	13.0	3025	10 ADE84158	Ade84158 Human lym
34	34	13.0	8711	4 AAS46699	Aas46699 Tumour su
35	34	13.0	13131	6 ABL92249	Ab192249 Chemicall
C 36	33.8	13.0	473	9 ACH28682	Ach28682 Human adu
37	33.8	13.0	562	5 ABV57017	Abv57017 Human pro
38	33.8	13.0	6545	10 ADE71194	Ade71194 Novel hum
39	33.8	13.0	37314	4 AAK71358	Aak71358 Human imm
C 40	33.8	13.0	172570	6 ABQ88207	Abq88207 Human ost
41	33.6	12.9	1974	4 AAS60993	Aas60993 Human can
42	33.6	12.9	2092	5 ABA15073	Abal5073 Human ner
C 43	33.6	12.9	3479	4 ABL03600	Ab103600 Drosophil
C 44	33.6	12.9	3968	4 ABL03602	Ab103602 Drosophil
45	33.6	12.9	5349	6 ABL32340	Ab132340 Human imm

ALIGNMENTS

RESULT 1
AEA49636
ID AEA49636 standard; cDNA; 261 BP.

XX AEA49636;

XX 11-AUG-2005 (first entry)

XX Loblolly pine cDNA clone, LPZ-061, SEQ ID 131.

XX Plant; embryogenesis; plant breeding; paper; wood; ss.

XX Pinus taeda.

XX US2005125161-A1.

XX 09-JUN-2005.

XX 02-SEP-2003; 2003US-00651991.

XX 11-OCT-2000; 2000US-0239250P.

XX 12-JAN-2001; 2001US-0260882P.

XX 11-OCT-2001; 2001US-00973994.

XX (PAPE-) INST PAPER SCI & TECHNOLOGY INC.

XX Cairney J, Xu N;

XX WPI; 2005-417092/42.

XX Relational database of cDNA molecules including those corresponding to

XX Loblolly pine major intrinsic protein, being differentially expressed

XX during plant embryogenesis, useful for staging plant embryos.

XX Claim 17; SEQ ID NO 131; 219pp; English.

XX The present invention relates to a relational database of cDNA molecules

XX comprising multiple nucleotide sequences (AEA49506-AEA49832). The

XX database is useful for staging plant embryos, selecting advantageous

XX plant clones, determining embryo fitness, and selecting advantageous

XX conditions for embryo development. The invention is particularly intended

CC for use in breeding trees with superior characteristics for use in
CC production of paper and other wood products.
XX
SQ Sequence 261 BP; 69 A; 43 C; 60 G; 89 T; 0 U; 0 Other;
Query Match 100.0%; Score 261; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.1e-63;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTGACCGTACAGCATTTATTGATGTTCTATTTTGTGTTGTTGCAAGTTTCCGATTTCG 60
DB 1 AGGTGACCGTACAGCATTTATTGATGTTCTATTTTGTGTTGTTGCAAGTTTCCGATTTCG 60
QY 61 CTGTGAGGCACGGAAACGAGATAAGTTGTAAGTTTGTCTCGCTGATTGAGGCACGGA 120
DB 61 CTGTGAGGCACGGAAACGAGATAAGTTGTAAGTTTGTCTCGCTGATTGAGGCACGGA 120
QY 121 AAACGAGATAAGTTGTAAATTTTGTCTCGCTGATTGTTGTTGCTGAATATTTCTCTCACTAT 180
DB 121 AAACGAGATAAGTTGTAAATTTTGTCTCGCTGATTGTTGTTGCTGAATATTTCTCTCACTAT 180
QY 181 AAAAGCATTTTCCAGAAATAAGAGAGCTTTTCGAACTGTTTCCCAAGAGTTGTAG 240
DB 181 AAAAGCATTTTCCAGAAATAAGAGAGCTTTTCGAACTGTTTCCCAAGAGTTGTAG 240
QY 241 GGGGTTTTTCCACGGTCACCT 261
DB 241 GGGGTTTTTCCACGGTCACCT 261

RESULT 2
ABQ67150
ID ABQ67150 standard; DNA; 40324 BP.
XX
AC ABQ67150;
XX
DT 28-AUG-2002 (first entry)
XX
DE Human angiogenesis associated polynucleotide SEQ ID NO 180.
XX
KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;
KW macular degeneration; inflammatory bowel disease; Crohn's disease;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiarteriosclerotic; ds.
XX
OS Homo sapiens.
XX
PN WO200246454-A2.
XX
PD 13-JUN-2002.
XX
PF 06-DEC-2001; 2001WO-EP014320.
XX
PR 06-DEC-2000; 2000DE-01061338.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Schacht O;
XX
DR WPI; 2002-500450/53.
XX
PT New nucleic acid fragments from chemically treated angiogenesis-
PT associated genes, useful for determining methylation status, e.g. in
PT diagnosis or treatment of cancer.
XX
PS Claim 1; SEQ ID NO 180; 41pp + Sequence Listing; German.
XX
CC The invention relates to a nucleic acid (I) comprising a segment of 18
CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
CC related oligomers, are used to evaluate the methylation status and/or
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for

CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 40324 BP; 12266 A; 436 C; 7652 G; 19970 T; 0 U; 0 Other;
Query Match 14.3%; Score 37.4; DB 6; Length 40324;
Best Local Similarity 52.2%; Pred. No. 4.7;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 25 TGTTCATTTTGTGTTGCAAGTTTTCGATTTCGCTGAGGCACGGAAACGAGATA 84
DB 26359 TGTTCATTTTGTGTTGCAAGTTTTCGATTTCGCTGAGGCACGGAAACGAGATA 26418
QY 85 AGTTGTAAAGTTTGTGCTGCTGATTGAGGCACGGAAACGAGATAAGTTGTAAGTTT 144
DB 26419 TTTTAAAGATTGTTGAGATGATTGATTTTAGATAATAGTTAGTATTAATTAGTT 26478
QY 145 GCTCGCTGATTTTGTGCTGAATATTTCTCTCACTATAAA 183
DB 26479 AATTGAAGATTTTATTGAGAAATAGTTTGTGTATATATA 26517
RESULT 3
AAC56859
ID AAC56859 standard; DNA; 357 BP.
XX
AC AAC56859;
XX
DT 25-JAN-2001 (first entry)
XX
DE Pinus radiata transcription factor DNA sequence #305.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS Pinus radiata.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US006112.
XX
PR 11-MAR-1999; 99US-00266513.
PR 18-AUG-1999; 99US-0149485P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI; 2000-579369/54.
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide.
XX
PS Claim 1; Page 531; 747pp; English
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.

Db 563 GATAAGTTGTAAGAAGTTTGCTTGCTGATTTTTTTTGCTGA-ATATTTCTCACTATAAAAAAG 505

QY 187 CATTTTCCAGAAATAAGAAGGAGCTTTCGAACTGGTTTTCCTCCCAAGAGTTGTAGGGGTTT 246
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 504 CATTTT-CAGAAATAAGAAGGAGCTCTCGAACTGGTTTCTCCAAGAGTTGTAGGGGTTT 446

QY 247 TTTCCACGGTCACCT 261

Db 445 TTTTCAGGGTCACCT 431

RESULT 2

DR053876/c

LOCUS RTCA1_13_H01.g1_A029 Roots minus calcium Pinus taeda cDNA clone

DEFINITION RTCA1_13_H01_A029 5', mRNA sequence.

ACCESSION DR053876

VERSION DR053876.1 GI:66977443

KEYWORDS EST.

SOURCE Pinus taeda (loblolly pine)

ORGANISM Pinus taeda

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 674)

AUTHORS Pratt,L., Cordonnier-Pratt,M.M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.

TITLE An EST database from calcium-deficient loblolly pine (Pinus taeda) roots

JOURNAL Unpublished (2005)

COMMENT Other_ESTs: RTCA1_13_H01.b1_A029
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES

source

1. .674

Location/Qualifiers

/organism="Pinus taeda"

/mol_type="mRNA"

/strain="3 CCLONES"

/db_xref="taxon:3352"

/clone="RTCA1_13_H01_A029"

/lab_host="DH10B-T1 phage-resistant E. coli"

/clone_lib="Roots minus calcium"

/note="Organ: Root; Vector: pSL1180; Site_1: EcoRI; Site_2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 139 days (July 28 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Starting five weeks prior to harvesting roots for mRNA preparation, the weekly fertilizer regimen for the potted trees was modified to omit calcium from the 0.5x Hoagland's solution (5 applications). For two days prior to starting the modified fertilizer regimen, pots were flushed extensively with water to reduce residual levels of calcium. Double-stranded cDNA was cloned

ORIGIN

Query Match 34.4%; Score 89.8; DB 9; Length 674;
Best Local Similarity 83.2%; Pred. No. 3.9e-13;
Matches 114; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 125 GAGATAAGTTGTAAAAATTTTGCTGCTGATTTTTTGTGAATATTTCTCACTATAAAA 184
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 511 GAGATAAGTTGTAAAAAGTTTGCTGCTGA-TTTTGTGTAATATTTCTTGTCTGTAAAAA 453

QY 185 AGCATTTTCCAGAAATAAGAAGAGCTTTCGAACCTGGTTTCCCAAGAGTTGTAGGGG 244
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 452 GACATTTTTCAGAAATAAAGAAGAGCTCTCAAACTGGTTTCCCAAGAGTTGTAGGGG 393

QY 245 TTTTTCACGGTCACCT 261

Db 392 TATTTCAGGGTCACAT 376

RESULT 3

DR120385

LOCUS RTMG1_29_H05_A029 Roots minus magnesium Pinus taeda cDNA clone

DEFINITION RTMG1_29_H05_A029 3', mRNA sequence.

ACCESSION DR120385

VERSION DR120385.1 GI:67708695

KEYWORDS EST.

SOURCE Pinus taeda (loblolly pine)

ORGANISM Pinus taeda

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 817)

AUTHORS Pratt,L., Cordonnier-Pratt,M.M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.

TITLE An EST database from magnesium-deficient loblolly pine (Pinus taeda) roots

JOURNAL Unpublished (2005)

COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: M13-21 (TGTAAGACGACGCCAGT)

FEATURES

source

1. .817

Location/Qualifiers

/organism="Pinus taeda"

/mol_type="mRNA"

/strain="3 CCLONES"

/db_xref="taxon:3352"

/clone="RTMG1_29_H05_A029"

/lab_host="DH10B-T1 phage-resistant E. coli"

/clone_lib="Roots minus magnesium"

/note="Organ: Root; Vector: pSL1180; Site_1: EcoRI; Site_2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 139 days (July 28 2003 harvest) under ambient conditions in a local

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:02:26 ; Search time 1779.21 Seconds
(without alignments)
9380.717 Million cell updates/sec

Title: US-09-973-994-131
Perfect score: 261
Sequence: 1 aggtgaccgtacagcattta.....gggtttttccacggtcacct 261

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*			
1:	gb_env:	*	
2:	gb_pat:	*	
3:	gb_ph:	*	
4:	gb_pl:	*	
5:	gb_pr:	*	
6:	gb_ro:	*	
7:	gb_sts:	*	
8:	gb_sy:	*	
9:	gb_un:	*	
10:	gb_vi:	*	
11:	gb_ov:	*	
12:	gb_htg:	*	
13:	gb_in:	*	
14:	gb_om:	*	
15:	gb_ba:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	42	16.1	171009	12 AC128851	AC128851 Rattus no
2	42	16.1	264832	12 AC106309	AC106309 Rattus no
C 3	41.8	16.0	38744	13 CEC56A3	Z77655 Caenorhabdi
4	41.2	15.8	161505	6 AC154234	AC154234 Mus muscu
5	41.2	15.8	222702	12 CT009561	CT009561 Mus muscu
6	40.8	15.6	182957	12 CR931799	CR931799 Danio rer
7	40.8	15.6	259146	12 CT573231	CT573231 Danio rer
C 8	40.6	15.6	231271	12 AC117004	AC117004 Rattus no
C 9	40.6	15.6	232508	12 AC116189	AC116189 Rattus no
C 10	39.8	15.2	237705	12 AC171299	AC171299 Bos tauru
C 11	39.8	15.2	239162	12 AC175441	AC175441 Bos tauru
C 12	39.4	15.1	94646	12 AC171129	AC171129 Helobdell
C 13	39.2	15.0	158135	12 AC080040	AC080040 Homo sapi
14	39.2	15.0	174058	12 AC068695	AC068695 Homo sapi
C 15	39.2	15.0	176570	5 AC009499	AC009499 Homo sapi
16	38.8	14.9	33651	13 AC114261	AC114261 Dictyoste
17	38.8	14.9	195124	11 CR548625	CR548625 Zebrafish
18	38.8	14.9	208436	12 CT027756	CT027756 Danio rer

19	38.6	14.8	7942	10 HPV52	X74481 Human papil
C 20	38.6	14.8	155399	5 PT027015	CR937025 Pan trogl
C 21	38.6	14.8	192290	12 AC153218	AC153218 Bos tauru
C 22	38.6	14.8	195220	12 AC151020	AC151020 Callithri
23	38.6	14.8	249791	11 BX537133	BX537133 Zebrafish
24	38.6	14.8	271777	12 AC157148	AC157148 Bos tauru
25	38.2	14.6	103409	11 AL662880	AL662880 Zebrafish
26	38.2	14.6	163804	11 BX005463	BX005463 Zebrafish
C 27	38.2	14.6	170688	6 AC101795	AC101795 Mus muscu
C 28	38.2	14.6	222137	12 CR954958	CR954958 Danio rer
C 29	38.2	14.6	281676	12 AC171784	AC171784 Bos tauru
30	38	14.6	35850	13 AC006791	AC006791 Caenorhab
C 31	38	14.6	230810	12 AC118784	AC118784 Rattus no
C 32	37.8	14.5	180309	5 AC073576	AC073576 Homo sapi
33	37.8	14.5	188107	5 CNS01DWE	AL137129 Human chr
34	37.8	14.5	264773	12 AC163148	AC163148 Bos tauru
35	37.6	14.4	73250	5 AC114946	AC114946 Homo sapi
36	37.6	14.4	114214	12 AP007470	AP007470 Lotus cor
37	37.6	14.4	168419	12 AC025384	AC025384 Homo sapi
C 38	37.6	14.4	224828	12 AC157113	AC157113 Bos tauru
39	37.4	14.3	40324	2 AX458634	AX458634 Sequence
C 40	37.4	14.3	110000	12 PFMA18P1_04	Continuation (5 of
41	37.4	14.3	125632	4 AC015446	AC015446 Arabidops
42	37.4	14.3	146130	12 AC015875	AC015875 Homo sapi
43	37.4	14.3	200195	5 AC146220	AC146220 Pan trogl
C 44	37.2	14.3	27426	5 AC079740	AC079740 Homo sapi
C 45	37.2	14.3	132822	12 AC181755	AC181755 Strongylo

ALIGNMENTS

RESULT 1
AC128851/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC128851 171009 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH30-445E24, WORKING DRAFT SEQUENCE.

AC128851 GI:25139471

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

1 (bases 1 to 171009)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Hulysk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE: Weinstock, G: and
Direct Submission

Direct submission
Unpublished

2 (pages 1 to 171009)

World K C
T 9991 2

Worley, K.L.
Direct Submission

Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 Bayliff Plaza, Houston,
3 (bases 1 to 171009)

3 (Bases 1 to 171003)
Pat Genome Sequencing Consortium

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23914598. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center: paytor C
Center code: BCM

Center code: BCM
Web site: <http://www.hqsc.bcm.tmc.edu/>

Contact: helena@wml.wisc.edu
 Web Site: <http://www.wml.wisc.edu>

----- Project Information -----
Contact: Ingsa-NetProc@mc:edu

----- PROJECT INFORMATION

Center project name: KAVY
Center clone name: CH230-445E24

Center clone name: CH230-445E

```
----- Summary Statistics
n      = 1000000
mean    = 0.9999999999999999
variance = 0.9999999999999999
```

Assembly program: Phrap; version 0.990329

Consensus quality: 164467 bases at least Q40

Consensus quality: 166274 bases at least Q30

Consensus quality: 167295 bases at least Q20

Estimated insert size: 171470; sum-of-contigs estimation

Quality co

```

-----
* NOTE: Estimated insert size may differ from sequence length
*       (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
*
* NOTE: This is a 'working draft' sequence. It currently
*       consists of 1 contigs. Gaps between the contigs
*       are represented as runs of N. The order of the pieces
*       is believed to be correct as given, however the sizes
*       of the gaps between them are based on estimates that have
*       provided by the submitter.
*
* This sequence will be replaced

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 11, 2006, 07:21:26 ; Search time 125.935 Seconds
(without alignments)
4561.330 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtgcgacctaagatgca.....gccatcactggatgcacc 307

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				Issued Patents NA:*	
	1:	/EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*			
	2:	/EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*			
	3:	/EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*			
	4:	/EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*			
	5:	/EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*			
	6:	/EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*			
	7:	/EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*			
	8:	/EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*			
	9:	/EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*			
	10:	/EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	
1	55.8	18.2	7218	2	US-08-232-463-14 Sequence 14, Appli
2	36.4	11.9	63588	3	US-09-873-404-3 Sequence 3, Appli
3	36.4	11.9	63588	3	US-10-243-735-3 Sequence 3, Appli
4	36.4	11.9	63588	4	US-10-730-010-3 Sequence 3, Appli
5	36	11.7	219964	3	US-09-949-016-15086 Sequence 15086, A
6	35.4	11.5	101349	3	US-09-949-016-17433 Sequence 17433, A
7	34.8	11.3	1664976	3	US-08-916-421B-1 Sequence 1, Appli
8	34.8	11.3	1664976	3	US-09-692-570-1 Sequence 1, Appli
9	34.6	11.3	5394	3	US-08-688-376-1 Sequence 1, Appli
10	34.4	11.2	764	3	US-09-949-016-5644 Sequence 5644, Ap
11	34.4	11.2	764	3	US-09-949-016-5645 Sequence 5645, Ap
12	34.4	11.2	1127	3	US-09-976-594-1102 Sequence 1102, Ap
13	34	11.1	15756	3	US-09-949-016-13607 Sequence 13607, A
14	34	11.1	29165	3	US-09-949-016-12340 Sequence 12340, A
15	33.8	11.0	1141	3	US-09-806-708B-22 Sequence 22, Appli
16	33.4	10.9	453	3	US-09-328-352-2667 Sequence 2667, Ap
17	33.4	10.9	601	3	US-09-949-016-64949 Sequence 64949, A
18	33.4	10.9	87863	3	US-09-949-016-14402 Sequence 14402, A
19	33.4	10.9	300402	3	US-09-949-016-13632 Sequence 13632, A
20	33	10.7	396	3	US-09-495-050A-98 Sequence 98, Appli
21	33	10.7	6854	2	US-08-468-036-4 Sequence 4, Appli
22	33	10.7	6854	2	US-08-376-843-4 Sequence 4, Appli
23	33	10.7	24740	3	US-09-949-016-13528 Sequence 13528, A

C	24	33	10.7	74962	3	US-09-685-853A-3	Sequence 3, Appli
	25	32.6	10.6	1377	3	US-09-583-110-2506	Sequence 2506, Ap
	26	32.6	10.6	1392	3	US-09-107-433-903	Sequence 903, App
	27	32.6	10.6	39318	3	US-09-949-016-13798	Sequence 13798, A
C	28	32.2	10.5	601	3	US-09-949-016-64948	Sequence 64948, A
C	29	32.2	10.5	601	3	US-09-949-016-152910	Sequence 152910, A
C	30	32.2	10.5	166698	3	US-09-949-016-16038	Sequence 16038, A
	31	32	10.4	601	3	US-09-949-002-3310	Sequence 3310, Ap
	32	32	10.4	601	3	US-09-949-002-3311	Sequence 3311, Ap
	33	32	10.4	601	3	US-09-949-002-7769	Sequence 7769, Ap
	34	32	10.4	601	3	US-09-949-002-7770	Sequence 7770, Ap
	35	32	10.4	51905	3	US-09-949-002-667	Sequence 667, App
	36	32	10.4	51905	3	US-09-949-002-781	Sequence 781, App
	37	31.8	10.4	228	3	US-09-248-796A-11240	Sequence 11240, A
C	38	31.8	10.4	601	3	US-09-949-016-152908	Sequence 152908, A
C	39	31.8	10.4	601	3	US-09-949-016-152909	Sequence 152909, A
C	40	31.8	10.4	1827	3	US-09-270-767-1308	Sequence 1308, Ap
C	41	31.8	10.4	1827	3	US-09-270-767-16590	Sequence 16590, A
	42	31.8	10.4	250352	3	US-09-949-016-14724	Sequence 14724, A
	43	31.6	10.3	100848	3	US-09-596-002-39	Sequence 39, Appli
C	44	31.6	10.3	312474	3	US-09-949-016-17434	Sequence 17434, A
C	45	31.4	10.2	601	3	US-09-949-016-128705	Sequence 128705, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:46:32 ; Search time 916.676 Seconds
(without alignments)
4115.199 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtggatcctagaattgca.....gccatcactcgatgcacc 307

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	307	10	US-10-651-991-79 Sequence 79, Appl
C 2	296	96.4	308	10	US-10-651-991-78 Sequence 78, Appl
C 3	36.6	11.9	165156	10	US-10-995-561-13304 Sequence 13304, A
4	36.4	11.9	63588	6	US-10-243-735-3 Sequence 3, Appli
5	36.4	11.9	63588	8	US-10-730-010-3 Sequence 3, Appli
6	36	11.7	511	4	US-09-925-065A-261418 Sequence 261418,
7	36	11.7	511	5	US-09-925-065A-261418 Sequence 261418,
8	36	11.7	548	12	US-10-301-480-339709 Sequence 339709,
9	36	11.7	548	12	US-10-301-480-953118 Sequence 953118,
10	36	11.7	661	3	US-09-998-598-965 Sequence 965, App
C 11	35.4	11.5	5976	8	US-10-437-963-8075 Sequence 8075, Ap
C 12	35.4	11.5	165156	8	US-10-741-601-5668 Sequence 5668, Ap
13	35.2	11.5	669	12	US-10-301-480-545653 Sequence 545653,
14	35.2	11.5	669	12	US-10-301-480-1159062 Sequence 1159062,
15	35	11.4	445	3	US-09-920-300A-620 Sequence 620, App
16	35	11.4	445	6	US-10-033-528-620 Sequence 620, App
17	35	11.4	445	7	US-10-099-926-620 Sequence 620, App

18	35	11.4	445	10	US-10-961-527-620 Sequence 620, App
19	35	11.4	18820	3	US-09-764-891-6452 Sequence 6452, Ap
20	35	11.4	18820	6	US-10-205-428-733 Sequence 733, App
C 21	34.6	11.3	793	12	US-10-301-480-561365 Sequence 561365,
C 22	34.6	11.3	793	12	US-10-301-480-1174774 Sequence 1174774,
C 23	34.6	11.3	3673778	7	US-10-312-841-1 Sequence 1, Appli
C 24	34.4	11.2	257	8	US-10-424-599-94708 Sequence 94708, A
C 25	34.4	11.2	306	8	US-10-242-535A-21758 Sequence 21758, A
C 26	34.4	11.2	306	8	US-10-085-783A-21758 Sequence 21758, A
C 27	34.4	11.2	321	8	US-10-242-535A-12694 Sequence 12694, A
C 28	34.4	11.2	321	8	US-10-085-783A-12694 Sequence 12694, A
C 29	34.4	11.2	426	8	US-10-242-535A-39856 Sequence 39856, A
C 30	34.4	11.2	426	8	US-10-085-783A-39856 Sequence 39856, A
C 31	34.4	11.2	436	8	US-10-242-535A-7614 Sequence 7614, Ap
C 32	34.4	11.2	436	8	US-10-085-783A-7614 Sequence 7614, Ap
C 33	34.4	11.2	482	8	US-10-242-535A-24125 Sequence 24125, A
C 34	34.4	11.2	482	8	US-10-085-783A-24125 Sequence 24125, A
C 35	34.4	11.2	533	8	US-10-242-535A-15691 Sequence 15691, A
C 36	34.4	11.2	533	8	US-10-085-783A-15691 Sequence 15691, A
C 37	34.4	11.2	540	16	US-10-128-061-3720 Sequence 3720, Ap
C 38	34.4	11.2	540	16	US-10-128-049-3720 Sequence 3720, Ap
39	34.4	11.2	541	4	US-09-925-065A-465048 Sequence 465048,
40	34.4	11.2	541	4	US-09-925-065A-465049 Sequence 465049,
41	34.4	11.2	541	5	US-09-925-065A-465048 Sequence 465048,
42	34.4	11.2	541	5	US-09-925-065A-465049 Sequence 465049,
43	34.4	11.2	544	12	US-10-301-480-521152 Sequence 521152,
44	34.4	11.2	544	12	US-10-301-480-521153 Sequence 521153,
45	34.4	11.2	544	12	US-10-301-480-1134561 Sequence 1134561,

ALIGNMENTS

RESULT 1

US-10-651-991-79
; Sequence 79, Application US/10651991
; Publication No. US20050125161A1
; GENERAL INFORMATION:
; APPLICANT: CAIRNEY, JOHN
; APPLICANT: XU, NANFIE
; TITLE OF INVENTION: DIFFERENTIALLY-EXPRESSED CONIFER CDNAs, AND THEIR USE
; TITLE OF INVENTION: IN IMPROVING SOMATIC EMBRYOGENESIS
; FILE REFERENCE: 7648.0023-00
; CURRENT APPLICATION NUMBER: US/10/651,991
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: 60/239,250
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/260,882
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 339
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 307
; TYPE: DNA
; ORGANISM: Pinus taeda
US-10-651-991-79

Query Match 100.0%; Score 307; DB 10; Length 307;
Best Local Similarity 100.0%; Pred No. 1.6e-81;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGTGGATCCTAGAAATTCATCTGTTGCCGTTGCTACTCCCTTTGAATTAGAACCTTCC	60
Db	1	GGTGGATCCTAGAAATTCATCTGTTGCCGTTGCTACTCCCTTTGAATTAGAACCTTCC	60
Qy	61	TTGCCACCTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTGGCATTATTCCTAGAAAT	120
Db	61	TTGCCACCTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTGGCATTATTCCTAGAAAT	120
Qy	121	CCTCTTGCTCCTTAACATCCTATAGGTCATCTCATGCCAAGAGATCTTGACAAATTGG	180
Db	121	CCTCTTGCTCCTTAACATCCTATAGGTCATCTCATGCCAAGAGATCTTGACAAATTGG	180

[illegible]

RESULT 2
US-10-517-441-108/c
; Sequence 108, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMRICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 108
; LENGTH: 9353
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-517-441-108

RESULT 3
US-10-501-187-222/c
; Sequence 222, Application US/10501187
; Publication No. US20060141455A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Rhonda
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS BREAST CELLS AND THEIR METHODS OF USE

```

; FILE REFERENCE: 2300-17767
; CURRENT APPLICATION NUMBER: US/10/501,187
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/345,637
; PRIOR FILING DATE: 2002-01-08
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-501-187-222

Query Match      11.2%; Score 34.4; DB 6; Length 583;
Best Local Similarity 59.0%; Pred. NO. 0.66;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      68  CTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTGGCATTATTCTTAGAATCCTCTTG 127
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      324  CAGGTTTTTTGGCTTTCTTTTGATTCATATTGTGCAAGACGTTCTTCCCTTAGCCCTCTTG 265
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      128  CTCCTTAACATCCTATAAGGTCATCTCATGCCAAGAGATC 167
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      264  CTTCTTCACCTTTCCCTCCTCATCATCAGATCCAAAGAGGTC 225
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4
US-11-266-748A-194452/c
; Sequence 194452, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIORITY FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 194452
; LENGTH: 809
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-194452

QY 128 CTCCTTAAACATCCTATAAGGTGTCATCTCATGCGCAAGAGATC 167

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:01:45 ; Search time 392.398 Seconds
(without alignments)
5454.872 Million cell updates/sec

Title: US-09-973-994-79
Perfect score: 307
Sequence: 1 ggtgcgacccctagaattgca.....gccatcactcggtcgacc 307

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_8:*
1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*
14:	geneseqn2005s:*
15:	geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	307	14	AEA49584 Loblolly
2	296	96.4	308	14	AEA49583 Loblolly
3	36.4	11.9	63588	8	ABS57150 Human gen
4	36.2	11.8	2000	11	ACL36306 Rice stre
5	36	11.7	661	6	ABV87654 Human col
6	35.8	11.7	677	1	AAN60393 Sequence
7	35.8	11.7	9353	13	ADS89445 Oligonuc1
8	35.6	11.6	265118	5	AAH41227 Pyrococcu
9	35.4	11.5	165156	13	ADS36459 Human aut
10	35	11.4	445	6	ABK45069 cDNA enco
11	35	11.4	2607	14	ADW16473 Eucalyptu
12	35	11.4	9353	13	ADS89092 Human AKR
13	35	11.4	18820	4	AAI03764 Human rep
14	35	11.4	18820	4	ABA07938 Human ova
15	34.8	11.3	64976	2	AAV21209_16 Continuation (17 o
16	34.6	11.3	5394	2	AAT00872 Murine mC
17	34.6	11.3	5394	3	AAZ86916 Mouse mC2
18	34.6	11.3	5394	14	AEC05449 Mouse mC2

C 19	34.4	11.2	521	10	ABZ84664	Abz84664 Toxicolog
C 20	34.4	11.2	540	14	AE814209	Ae814209 Hamster S
C 21	34.4	11.2	553	6	ABV89065	ABV89065 Human col
22	34.4	11.2	561	6	ABK44654	ABK44654 cDNA enco
23	34.4	11.2	563	6	ABK45341	ABK45341 cDNA enco
C 24	34.4	11.2	579	14	AE810567	Ae810567 Hamster c
C 25	34.4	11.2	583	10	ADK11816	Adk11816 Breast ca
C 26	34.4	11.2	583	14	ACL53983	ACL53983 Human col
C 27	34.4	11.2	619	6	ABK55781	ABK55781 Human col
C 28	34.4	11.2	741	6	ABQ55781	ABQ55781 Human ova
C 29	34.4	11.2	818	10	ADG32718	Adg32718 Human DNA
C 30	34.4	11.2	821	13	ACN37817	Acn37817 Tumour-as
C 31	34.4	11.2	821	14	ADZ48889	Adz48889 Insulin s
C 32	34.4	11.2	961	13	ACN37818	Acn37818 Tumour-as
C 33	34.4	11.2	961	14	ADZ70690	Adz70690 Human CDN
C 34	34.4	11.2	1089	4	AAK53287	Aak53287 Human pol
C 35	34.4	11.2	1126	4	AAK52303	Aak52303 Human pol
C 36	34.4	11.2	1127	12	ADL15373	Adl13373 Human ste
37	34.4	11.2	1147	4	AAD08441	Aad08441 Human sec
38	34.4	11.2	1147	5	AAD05439	Aad05439 Human sec
C 39	34.4	11.2	1155	5	AA85654	AA85654 DNA encod
C 40	34.4	11.2	1343	14	AED73250	Aed73250 Human pla
C 41	34.4	11.2	2453	13	ADX54786	Adx54786 Plant ful
C 42	34.4	11.2	80963	14	AEA61095_3	Continuation (4 of
C 43	34.2	11.1	9353	13	ADS89719	Ads89719 Oligonuc1
C 44	34	11.1	562	3	AAF16377	Aaf16377 Human pro
C 45	34	11.1	659	6	ABQ60765	ABQ60765 Human col

ALIGNMENTS

RESULT 1	
AEA49584	
ID	AEA49584 standard; cDNA; 307 BP
XX	
AC	AEA49584;
XX	
DT	11-AUG-2005 (first entry)
XX	
DE	Loblolly pine cDNA clone, LPS-0967 SEQ ID 79.
XX	
KW	Plant; embryogenesis; plant breeding; paper; wood; ss.
XX	
OS	Pinus taeda.
XX	
PN	US2005125161-A1.
XX	
PD	09-JUN-2005.
XX	
PF	02-SEP-2003; 2003US-00651991.
XX	
PR	11-OCT-2000; 2000US-0239250P.
PR	12-JAN-2001; 2001US-0260882P.
PR	11-OCT-2001; 2001US-00973994.
XX	
PA	(PAPE-) INST PAPER SCI & TECHNOLOGY INC.
XX	
PI	Cairney J, Xu N;
XX	
DR	WPI; 2005-417092/42.
XX	
PT	Relational database of cDNA molecules including those corresponding to
PT	Loblolly pine major intrinsic protein, being differentially expressed
PT	during plant embryogenesis, useful for staging plant embryos.
XX	
PS	Claim 17; SEQ ID NO 79; 219pp; English.
XX	
CC	The present invention relates to a relational database of cDNA molecules
CC	comprising multiple nucleotide sequences (AEA49506-AEA49832). The
CC	database is useful for staging plant embryos, selecting advantageous
CC	plant clones, determining embryo fitness, and selecting advantageous
CC	conditions for embryo development. The invention is particularly intended

CC for use in breeding trees with superior characteristics for use in
CC production of paper and other wood products.
XX
SQ Sequence 307 BP; 62 A; 84 C; 51 G; 110 T; 0 U; 0 Other;
Query Match 100.0%; Score 307; DB 14; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.2e-79;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTGCGATCCTAGAAATGCAATGCTGTTGCCGTTGCTACTCCCTTTGAATTAGAAGCCTTCC 60
Db 1 GGTGCGATCCTAGAAATGCAATGCTGTTGCCGTTGCTACTCCCTTTGAATTAGAAGCCTTCC 60
Qy 61 TTGCCACCTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTTGGCATTATTCCTAGAAT 120
Db 61 TTGCCACCTGGTCTTGAAGCATCCTTTTAGTCTTATTATCCTTTGGCATTATTCCTAGAAT 120
Qy 121 CCTCTTGCTCCTTAACATCCTATAAGGTCACTCATGCCCAAGAGATCTTGACAAATTTGG 180
Db 121 CCTCTTGCTCCTTAACATCCTATAAGGTCACTCATGCCCAAGAGATCTTGACAAATTTGG 180
Qy 181 AAGAATCTGCCAATTGTCTCCGCCCTCCTTAACCTGAAGTTGACTCTCTAAGTTGCTGAA 240
Db 181 AAGAATCTGCCAATTGTCTCCGCCCTCCTTAACCTGAAGTTGACTCTCTAAGTTGCTGAA 240
Qy 241 CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATATGCTTTTCTTTGTGCCATCACTCGGA 300
Db 241 CCTTCTGAGCTACCTGTTCTTTGTTTCTATCATATGCTTTTCTTTGTGCCATCACTCGGA 300
Qy 301 TCGCACC 307
Db 301 TCGCACC 307
RESULT 2
AEA49583/c
ID AEA49583 standard; cDNA; 308 BP.
XX
AC AEA49583;
XX
DT 11-AUG-2005 (first entry)
XX
DE Lobolly pine cDNA clone, LPS-095, SEQ ID 78.
XX
KW Plant; embryogenesis; plant breeding; paper; wood; ss.
XX
OS Pinus taeda.
XX
PN US2005125161-A1.
XX
PD 09-JUN-2005.
XX
PF 02-SEP-2003; 2003US-00651991.
XX
PR 11-OCT-2000; 2000US-0239250P.
PR 12-JAN-2001; 2001US-0260882P.
PR 11-OCT-2001; 2001US-00973994.
XX
XX (PAPE-) INST PAPER SCI & TECHNOLOGY INC.
XX
PI Cairney J, Xu N;
XX
XX WPI; 2005-417092/42.
DR
XX Relational database of cDNA molecules including those corresponding to
PT Lobolly pine major intrinsic protein, being differentially expressed
PT during plant embryogenesis, useful for staging plant embryos.
XX
PS Claim 17; SEQ ID NO 78; 219pp; English.
XX
CC The present invention relates to a relational database of cDNA molecules
CC comprising multiple nucleotide sequences (AEA49506-AEA49832). The
CC database is useful for staging plant embryos, selecting advantageous.

CC plant clones, determining embryo fitness, and selecting advantageous
CC conditions for embryo development. The invention is particularly intended
CC for use in breeding trees with superior characteristics for use in
CC production of paper and other wood products.
XX
SQ Sequence 308 BP; 110 A; 51 C; 84 G; 63 T; 0 U; 0 Other;
Query Match 96.4%; Score 296; DB 14; Length 308;
Best Local Similarity 99.7%; Pred. No. 2e-76;
Matches 307; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 GGTGCGATCCTAGAAATGCAATGCTGTTGCCGTTGCT-ACCTCCCTTTGAATTAGAAGCCTTC 59
Db 308 GGTGCGATCCTAGAAATGCAATGCTGTTGCCGTTGCTAATCCCTTTGAATTAGAAGCCTTC 249
Qy 60 CTTGCCACCTGGTCTTGAAGCATGCTTTTAGTCTTATTATCCTTGGCATTATTCCTAGAA 119
Db 248 CTTGCCACCTGGTCTTGAAGCATGCTTTTAGTCTTATTATCCTTGGCATTATTCCTAGAA 189
Qy 120 TCCTCTTGCTCCTTAACATCCTATAAGGTCACTCATGCCAAGAGATCTTGAATAATTTG 179
Db 188 TCCTCTTGCTCCTTAACATCCTATAAGGTCACTCATGCCAAGAGATCTTGAATAATTTG 129
Qy 180 GAAGAATCTGCCAATTGTCTCCGCCCTCCTTAACCTGAAGTTGACTCTTAAGTTGCTGA 239
Db 128 GAAGAATCTGCCAATTGTCTCCGCCCTCCTTAACCTGAAGTTGACTCTTAAGTTGCTGA 69
Qy 240 ACCTTCTGAGCTACCTGTTCTTTGTTTCTATCATGCTTTTCTTGCCATCACTCGG 299
Db 68 ACCTTCTGAGCTACCTGTTCTTTGTTTCTATCATGCTTTTCTTGCCATCACTCGG 9
Qy 300 ATCGCACC 307
Db 8 ATCGCACC 1
RESULT 3
ABS57150
ID ABS57150 standard; DNA; 63588 BP.
XX
AC ABS57150;
XX
DT 04-FEB-2003 (first entry)
XX
DE Human gene encoding a serine/threonine kinase.
XX
KW Human; ds; gene; serine/threonine kinase; lung carcinoma; chromosome 13;
KW SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key
FT variation Location/Qualifiers
FT replace(1848,G)
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1993,C)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT 2012. .60590
FT /*tag= a
FT /product= "Ser/Thr kinase"
FT 2012. .2128
FT /*tag= b
FT /number= 1
FT 2129. .10078
FT /*tag= c
FT /number= 1
FT replace(3796,T)
FT /*tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT replace(4854,G)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism"

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:10:56 ; Search time 3077.57 Seconds
(without alignments)
5578.184 Million cell updates/sec

Title: US-09-973-994-79

Perfect score: 307

Sequence: 1 ggtgcgactctagaattgca.....gccatcactcgatgcacc 307

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est3:*
- 3: gb_est4:*
- 4: gb_est5:*
- 5: gb_est6:*
- 6: gb_est7:*
- 7: gb_est8:*
- 8: gb_est9:*
- 9: gb_est10:*
- 10: gb_est11:*
- 11: gb_est12:*
- 12: gb_est13:*
- 13: gb_est14:*
- 14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.2	14.4	482	12	BZ833608 CH240_258
2	42.6	13.9	997	14	AL060767 Drosophil
3	42.4	13.8	466	1	AL794888 AL794888
4	42.4	13.8	862	10	DT441734 JGI_CABK3
5	41.4	13.5	738	14	BX159004 Danio rer
6	40.8	13.3	656	1	AL896255 AL896255
7	40	13.0	911	14	AG887137 Oryza sat
8	40	13.0	936	11	AZ209923 SP_0150_B
9	39.2	12.8	694	10	DT853440 LB0058.CR
10	39	12.7	791	10	DV553586 DC05004D0
11	38.8	12.6	532	4	CD056130 HO11E04S
12	38.8	12.6	617	8	CO010537 EST798872
13	38.6	12.6	763	5	CJ460131 CJ460131
14	38.6	12.6	808	9	CX444624 JGI_XZG10
15	38.6	12.6	939	14	DU958506 DU958506
16	38.4	12.5	549	10	W52992 zc02e10.r1
17	38.4	12.5	595	10	DR431991 DR431991
18	38.4	12.5	653	2	BJ636126 BJ636126
19	38.4	12.5	800	2	BJ643839 BJ643839

20	38.4	12.5	1017	11	AQ325859	AQ325859 nbxb0021H
21	38	12.4	728	13	CW262050	CW262050 104_729_1
22	38	12.4	753	11	BH452101	BH452101 BOHNU61TR
23	38	12.4	1101	10	DV012710	DV012710 CNB276-C0
24	37.8	12.3	440	2	BI449723	BI449723 dae74g01.
25	37.8	12.3	442	10	DT467317	DT467317 GH_CHX20D
26	37.8	12.3	791	10	DT465391	DT465391 GH_CHX17E
27	37.8	12.3	1101	14	CNS016UE	AL107216 Drosophil
28	37.6	12.2	485	1	AA782790	AA782790 aj08907.s
29	37.6	12.2	744	9	CX458263	CX458263 JGI_XZG27
30	37.6	12.2	765	9	CX413015	CX413015 JGI_XZT29
31	37.6	12.2	905	7	BF665751	BF665751 602124084
32	37.4	12.2	496	13	CL894723	CL894723 abg27h05.
33	37.4	12.2	850	10	DR898269	DR898269 JGI_XZT45
34	37.4	12.2	1101	14	CNS0182P	AL108811 Drosophil
35	37.4	12.2	1298	14	AG311035	AG311035 Mus muscu
36	37	12.1	814	14	BX995750	BX995750 Reverse s
37	37	12.1	893	1	AL553628	AL553628 AL553628
38	37	12.1	1213	2	BI560283	BI560283 603253520
39	36.8	12.0	835	11	BZ251414	BZ251414 CH230-262
40	36.8	12.0	901	13	CZ984248	CZ984248 198984 To
41	36.8	12.0	949	13	DU058624	DU058624 90441 Tom
42	36.6	11.9	763	14	CNS00410V	AL269608 Tetraodon
43	36.6	11.9	785	5	CF285981	CF285981 AGENCOURT
44	36.6	11.9	866	10	DV036164	DV036164 BRS1655 s
45	36.4	11.9	715	11	BH297774	BH297774 CH230-57H

ALIGNMENTS

RESULT 1

BZ833608

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BZ833608 482 bp DNA linear GSS 18-MAR-2003
CH240_258K11.TV CHORI-240 Bos taurus genomic clone CH240_258K11,
genomic survey sequence.
BZ833608
BZ833608.1 GI:29060400
GSS.
Bos taurus (cattle)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 482)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
Crawford,A.M. and McEwan,J.C.
Bovine BAC End Sequences from Library CHORI-240
Unpublished (2003)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBMC) by AgResearch Ltd., New Zealand and The
Institute of Genomic Research (TIGR), USA.

FEATURES

source

Location/Qualifiers
1..482
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 07:02:26 ; Search time 2092.79 Seconds
(without alignments)
9380.717 Million cell updates/sec

Title: US-09-973-994-79

Perfect score: 307

Sequence: 1 ggtgcatcctagaattgca.....gccatcactcgatcgacc 307

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:**

- 1: gb_env:**
- 2: gb_pat:**
- 3: gb_ph:**
- 4: gb_pl:**
- 5: gb_pr:**
- 6: gb_ro:**
- 7: gb_sts:**
- 8: gb_sy:**
- 9: gb_un:**
- 10: gb_vi:**
- 11: gb_ov:**
- 12: gb_htg:**
- 13: gb_in:**
- 14: gb_om:**
- 15: gb_ba:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.8	18.2	7218	2	I66494 Sequence 14
2	42.6	13.9	230556	12	AC162231 Bos tauru
3	42	13.7	110000	12	Continuation (5 of
4	40.4	13.2	230071	12	AC132699 Rattus no
5	40.4	13.2	338915	12	AC130523 Rattus no
6	39.8	13.0	45027	13	CBRG39N14
7	39	12.7	224084	12	CR318629
8	39	12.7	237446	11	BX546500 Zebrafish
9	38.8	12.6	110000	12	Continuation (2 of
10	38.6	12.6	7761	10	OPU83594
11	38.4	12.5	171562	5	AC107027
12	38.4	12.5	197206	12	AC023757
13	38.4	12.5	222975	12	AC130905
14	38.4	12.5	234072	12	AC099135
15	38.2	12.4	183481	12	AC120912
16	38.2	12.4	234844	12	AC111734
17	38	12.4	187571	12	AC177302
18	38	12.4	221478	12	AC079545

19	37.8	12.3	110000	4	CR382130 32	Continuation (33 o
20	37.8	12.3	116409	6	AC131118	AC131118 Mus muscu
21	37.8	12.3	153310	13	AC159433	AC159433 Trypanoso
22	37.8	12.3	156548	13	AC159413	AC159413 Trypanoso
23	37.8	12.3	206480	11	BX255911	BX255911 Zebrafish
24	37.6	12.2	184463	6	AC126553	AC126553 Mus muscu
25	37.6	12.2	216501	6	AC107795	AC107795 Mus muscu
26	37.6	12.2	219784	6	AC112792	AC112792 Mus muscu
27	37.4	12.2	156045	12	AC144885	AC144885 Papio anu
28	37.4	12.2	166583	12	AC119613	AC119613 Rattus no
29	37.4	12.2	166616	12	AC145761	AC145761 Papio anu
30	37.4	12.2	201602	12	AC074146	AC074146 Mus muscu
31	37.4	12.2	237286	12	AC113645	AC113645 Rattus no
32	37.2	12.1	110000	4	AP008218 088	Continuation (89 o
33	37.2	12.1	151176	4	CNS080CS	AL954154 Oryza sat
34	37.2	12.1	348034	13	CR382400	CR382400 Plasmodiu
35	37	12.1	52052	12	AC100081	AC100081 Mus muscu
36	37	12.1	123748	6	BX678774	BX678774 Mouse DNA
37	37	12.1	177857	12	AC135432	AC135432 Rattus no
38	37	12.1	202838	12	CR388230	CR388230 Mus muscu
39	36.8	12.0	85903	12	AC177439	AC177439 Strongylo
40	36.8	12.0	120077	4	AC148292	AC148292 Medicago
41	36.8	12.0	137281	12	AC124217	AC124217 Medicago
42	36.8	12.0	149635	11	BX510957	BX510957 Zebrafish
43	36.8	12.0	217726	6	AL928680	AL928680 Mouse DNA
44	36.8	12.0	254499	12	AC133758	AC133758 Rattus no
45	36.6	11.9	70932	12	AC174554	AC174554 Strongylo

ALIGNMENTS

RESULT 1	I66494	7218 bp	DNA	linear	PAT 28-DEC-1997
LOCUS	I66494				
DEFINITION	Sequence 14 from patent US 5670367.				
ACCESSION	I66494				
VERSION	I66494.1	GI:2724471			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 7218)				
AUTHORS	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.				
TITLE	Recombinant fowlpox virus				
JOURNAL	Patent: US 5670367-A 14 23 SEP-1997;				
FEATURES	Location/Qualifiers				
source	1..7218				
	/organism="unknown"				
	/mol_type="unassigned DNA"				

ORIGIN

Query Match	18.2%	Score: 55.8;	DB 2;	Length 7218;
Best Local Similarity	2.0%	Pred. No. 4.7e-07;		
Matches	6;	Conservative 186;	Mismatches 103;	Indels 0; Gaps 0;
Qy	3	TGCGATCCTAGAAATTCATCTGTTGCGGTTGCTACTCCCTTTGAATTAGAACCTTCCTT	62	
Db	1061	TGCGATYYY	1120	
Qy	63	GCCACCTGGTCTTGAAGCATCCTTTAGCTTATTATCCTTGGCATTATTCCTAGAAATCC	122	
Db	1121	YYY	1180	
Qy	123	TCTTGCTCCTTAACATCCTATAAGGTCATCTCATGCGAAGAGATCTTGACAAATTTGGAA	182	
Db	1181	YYY	1240	
Qy	183	GAATCTGCCAATTGCTCCCGCCCTCTTAACCTGAAGTGACTCTCTAAGTTGCTGAACC	242	
Db	1241	YYY	1300	
Qy	243	TTCTGAGCTACCTGTTCTTTTCTATCATGCTTTTCTTTTGTGTCATCACTC	297	

